Express Mailing Label No. EL 985216526 US

PATENT APPLICATION
Docket No. SLA1196

UNITED STATES PATENT APPLICATION

of

JOHN E. DOLAN

AND

JON M. SPEIGLE

for

SYSTEMS AND METHODS FOR ILLUMINANT MODEL ESTIMATION

SYSTEMS AND METHODS FOR ILLUMINANT MODEL ESTIMATION

RELATED CASES

The present application is a continuation-in-part of the following three applications:

U.S. Patent Application Serial No. 10/677,034, entitled Systems and Methods for Computing the Presence of Self-Luminous Elements in an Image, invented by Jon M. Speigle and John E. Dolan, filed on September 30, 2003;

U.S. Patent Application Serial No. 10/676,306, entitled Systems and Methods for Correcting Image Color Balance and, invented by Jon M. Speigle and John E. Dolan, filed on September 30, 2003; and

U.S. Patent Application Serial No. 10/677,009, entitled Systems and Methods Illuminant Estimation, invented by John E. Dolan and Jon M. Speigle, filed on September 30, 2003.

TECHNICAL FIELD

[01] The present invention relates generally to digital image processing and more particularly to methods and systems for estimation of an image illuminant model.

BACKGROUND

[02] Colors viewed in an image are dependent on the light that illuminates the subject of the image. Different illuminants will lead to differences in reflected light from the surfaces of the image subject matter. The human visual system approximately corrects these differences in reflected light so that perceived surface color is approximately constant. However, when images are captured on media and viewed under a light source different than the source in the image scene, these natural corrections do not take place. Accordingly, it may be preferable for recorded images to be color-balanced to a standard, reference light source in order to appear as they would to the natural eye. This balancing or color correction can be performed once the scene illuminant is identified.

[03] Under known methods, illuminant estimation employs a fixed set of known illuminants, which are each characterized by the gamut of chromaticities that are possible under that illuminant. A chromaticity histogram may be computed for the image and compared with model histograms

SUMMARY

[04] Embodiments of the present invention provide methods and systems for image illuminant estimation by model matching with distance metrics.

BRIEF DESCRIPTION OF THE DRAWINGS

[05] The present embodiments will become more fully apparent from the following description and appended claims, taken in conjunction with the accompanying drawings. Understanding that these drawings depict only typical embodiments and are, therefore, not to be considered limiting of the invention's scope, the embodiments will be described with additional specificity and detail through use of the accompanying drawings in which:

Fig. 1 is a diagram illustrating a set of candidate illuminants as x-y chromaticity coordinates;

Fig. 2 is a diagram of an exemplary ordering scheme; and

Fig. 3 is a diagram of exemplary match surfaces for a given image and the illuminant grid of Figure 1.

DETAILED DESCRIPTION

[06] Illuminant estimation may be approached through a model matching strategy. In such a regime, a fixed set of illuminants is modeled. Modeling may be performed parametrically, by sample statistics or by other methods. The model that best accounts for the image data is chosen as the scene illuminant most likely to have produced the image. This decision process relies on computing a similar parametric or statistical description from the image and then performing a matching procedure of the image description with respect to the model base in order to select the "best" model.

[07] An exemplary model set, shown in Figure 1, consists of 81 illuminants plotted in x-y chromaticity space. Chromaticity coordinates are shown as dots 2. In this particular example, these coordinates have been regularly sampled in CIE-Lab coordinates centered on the D65 white point 4, which is a daylight reference illuminant. The coordinates are then mapped to x-y chromaticity space as displayed in Figure 1. In a similar fashion, the pixels of the input image may be transformed to CIE-xyY space (which is the x-y chromaticity space together with luminance information).

[08] In other embodiments, color values may be represented in an alternative colorspace or other chromaticity space such as

$$r = R/(R+G+B)$$

$$g = G/(R+G+B)$$

$$b = B/(R+G+B)$$

[09] In some embodiments of the present invention a distance metric may be used to determine which model best matches the image illuminant. The term "distance" is usually used to describe a length in space, such as the length of a line segment in two- or three-dimensional space. However, in this document, the term distance may also refer to a similar metric in a space with more than three dimensions. The term distance may also refer to a difference between two values that can be represented as points in multi-dimensional space.

[10] A measure d(x, y) is a distance metric iff (if and only if) it satisfies the following 3 conditions:

1. $d(x, y) = 0$ iff $x = y$ [zero only for equality]

2. d(x, y) = d(y, x) [symmetry]
3. d(x, y) + d(y, z) <= d(x, z) [triangle inequality]

It is usual (but often implicit) to add the condition of non-negativity—i.e., d(x, y) >= 0 for all x, y.

[11] A distance metric may also be referred to as a difference metric or a dissimilarity metric. In this specification and claims the terms distance, difference and dissimilarity are synonymous.

[12] Some embodiments of the present invention may perform operations on a subset of the full model set. The full model set may be pared down to a subset through an initial selection process. Final illuminant selection will then take place within the subset of models.

[13] In other embodiments, evaluations can also be performed in *greedy* fashion. The following is an example of greedy evaluation:

a) Evaluate a first best guess model as the prior value
b) Evaluate a second best guess model as the current value
c) Compute the difference vector between the current and prior values
d) While the difference vector is decreasing and greater than a criterion magnitude
   i) Set the prior value to the minimum of prior and current values
   ii) Evaluate a new best guess as the current value. Typically, the difference vector is used to guide the choice of the new model.
   iii) Compute the difference in scores between the prior and current values

[14] In embodiments of the present invention, it is possible to use any element-by-element vector distance/difference/dissimilarity measure. Such measures are appropriate for vectors or distributions on the same space.

[15] In some embodiments, distances under Minkowski norms such as L1, L2, L∞ may be used as a metric. These parameters may be expressed in equation form as follows:

$$\mathrm{d}_{\mathrm{L1}}(x,y)=\sum|x_i-y_i|$$
$$\mathrm{d}_{\mathrm{L2}}(x,y)=\left(\sum(x_i-y_i)^2\right)^{1/2}$$
$$\mathrm{d}_{\mathrm{L}\infty}(x,y)=\max|x_i-y_i|$$

[16] In other embodiments of the present invention, a Chi-squared statistical difference may be used as a metric. This difference metric may be expressed in equation form as follows:

$$d_\chi(x,y) = 2\chi^2 = \sum \begin{cases} \frac{(x_i - y_i)^2}{x_i + y_i} & : x_i + y_i \neq 0 \\ 0 & : x_i + y_i = 0 \end{cases}$$

[17] In still other embodiments a Jeffrey divergence metric may be used. This metric may be expressed in equation form as follows:

$$d_J(x,y) = \sum \begin{cases} x_i \log \frac{2x_i}{x_i + y_i} + y_i \log \frac{2y_i}{x_i + y_i} & : x_i + y_i \neq 0,\ x_i \cdot y_i \neq 0 \\ x_i \log 2 & : x_i \neq 0,\ y_i = 0 \\ y_i \log 2 & : x_i = 0,\ y_i \neq 0 \\ 0 & : x_i + y_i = 0 \end{cases}$$

Because of the log function, this measure is appropriate only when the vector elements are positive (i.e., $x_i \geq 0$). This non-negative condition holds for frequency vectors such as color histograms.

[18] In still other embodiments of the present invention, it is also possible to use non element-by-element distance/difference/dissimilarity measures. With such measures there may be a cost associated with matching non-corresponding elements—sometimes referred to as the *ground distance*—and this cost is added to the difference computed between the non-corresponding elements. Thus, the distance between two arbitrary elements of two vectors $x_i \in X$ and $y_j \in Y$ depends both on the functional difference in the elements (e.g., $\|x_i - y_j\|$) and the difference between *i* and *j*.

[19] In some of these embodiments, a quadratic form metric can be used. This metric can be expressed in equation form as follows:

$$d_Q(x,y) = \sqrt{(x-y)^T \mathbf{Q}(x-y)}$$

where the quadratic form is a similarity transform—i.e., $\mathbf{Q} = \mathbf{Q}^\mathrm{T}$ and $\mathbf{Q}^\mathrm{T}\mathbf{Q} = \mathbf{Q}\mathbf{Q}^\mathrm{T} = \mathbf{I}$. Elements of $\mathbf{Q} = \lfloor q_j \rfloor$ encode the normalized ground distances between elements: $q_{ij} = 1 - d_{ij}/d_{\max}$. Note that in the case that $\mathbf{Q} = \mathbf{I}$, this distance reduces to the Euclidean (L2) distance.

[20] In still further embodiments, a cumulative match distance may be used. This metric may be expressed in equation form as follows:

$$d_M(x,y) = \sum |\hat{x}_i - \hat{y}_i|$$

where $\hat{x}_i = \sum_{j \le i} x_j$ denotes the cumulative histogram of x up to cell i and similarly for y. Note that this can generalize to multidimensional distributions as follows:

$$x_{mnp} = \sum_{i \le m} \sum_{j \le n} \sum_{k \le p} |x_{ijk} - y_{ijk}|.$$

[21] In other embodiments a projective match distance may be used as a metric. This metric may be expressed in equation form as follows:

$$d_{PM}(x,y) = \sum |\tilde{x}_i - \tilde{y}_i|.$$

[22] Here $\tilde{x}_i = \sum_{j \le i} x_j$ is the cumulative sum of x up to cell i after projecting the possibly multidimensional histogram to a one dimensional structure—i.e., by imposing a total ordering on the histogram bins. Strictly speaking this is NOT the cumulative distribution, except in the case that the histogram is originally one dimensional. Ordering schemes include distance (L1, L2, …) from the origin, zig-zag, row or column major, etc. Some examples are shown in Figure 2. Note that this measure may be an approximation to the cumulative distance (discussed above) and one that is more efficient to compute. It is important and obvious that the same ordering should be imposed on both histograms.

[23] Additional illuminant estimation complexity occurs when an image contains non-reflective or self-luminous objects that do not directly reflect light from an illuminant. Conventional correction algorithms assume that all image pixels represent reflecting surfaces. When an image contains self-luminous objects such as sky and other light sources the surface-pixel assumption is violated. When an image contains a significant portion of non-reflective, self-luminous objects, conventional methods will fail and the image illuminant will be incorrectly determined. For example, if an image contains blue sky and the color-balance algorithm assumes that all pixels are reflecting objects, "bluish" pixels could be taken as evidence that the illumination of the scene is bluish. Because a color correction is approximately the opposite hue of the estimated illuminant, the correction for a bluish illuminant would be to

shift the image in a yellowish direction. This correction might produce an overly yellowish ground/surface region and a desaturated sky region.

[24] These color correction, color balance or color constancy algorithms generally do not address the question of how to handle images containing luminous objects, which are also referred to herein as self-luminous objects. They have, rather, focused on images where the surface-pixel assumption is satisfied (e.g., uniformly illuminated Mondrian-like images).

[25] In some embodiments of the present invention, for each pixel, image element or image area, a function is computed that estimates the probability p that the pixel, element or area corresponds to a reflective surface—i.e., the probability that it is not self-luminous. This is necessary because nearly all illuminant estimation procedures assume that the image pixels are a product of the illuminant reflecting from surfaces in the scene. So, incorporating self-luminous pixels would invalidate the estimate. Note that this computation is unnecessary for the model histograms, because these histograms are by definition constructed only from data that corresponds to reflective surfaces.

[26] The value computed by this function may be used to increment the accumulator of the corresponding histogram bin $(\bar{x},\bar{y})$, and the function has the following form:

$$\varphi(x,y,Y,r,c) \mapsto (\bar{x},\bar{y},p).$$

Some of these functions may be defined on a 5-tuple consisting of 2 chromaticity values x, y; 1 luminance value Y; and 2 image position values r, c, corresponding to row and column positions. It may return a 3-tuple consisting of the corresponding histogram bin indices $(\bar{x},\bar{y})$ and the bin increment, p. (Note that, in some embodiments, p ranges on the interval (0..1); it will be 0 for pixels corresponding to non-reflective objects and 1 only for pixels corresponding to reflective surfaces.)

[27] Details of embodiments of these functions are explained in detail in separate, previously-filed patent applications:

U.S. Patent Application Serial No. 10/677,034, entitled Systems and Methods for Computing the Presence of Self-Luminous Elements in an Image, invented by Jon M. Speigle and John E. Dolan, filed on September 30, 2003;

U.S. Patent Application Serial No. 10/676,306, entitled Systems and Methods for Correcting Image Color Balance and, invented by Jon M. Speigle and John E. Dolan, filed on September 30, 2003; and

U.S. Patent Application Serial No. 10/677,009, entitled Systems and Methods Illuminant Estimation, invented by John E. Dolan and Jon M. Speigle, filed on September 30, 2003., which are hereby incorporated by reference into this application.

[28] In other embodiments where the image color histogram is computed over three or more dimensions, the function estimating the probability that the pixel element or area corresponds to a reflective surface would accept arguments of the color coordinates and pixel coordinates.

[29] In other embodiments the function estimating probability of an element or area corresponding to a reflective surface would operate on a neighborhood of image pixel elements.

[30] In some of these embodiments, a histogram encodes the amount of evidence in the image for the presence of each chromaticity originating from a reflective surface.

[31] In some embodiments of the present invention, the image histogram is matched to the model histograms and the best match illuminant is selected. The match metric used in some embodiments is the chi-squared statistic, which is a procedure to test the variance of a sample population against an underlying population. In embodiments of the present invention, the image may represent a sample population and an illuminant model may be represented as the hypothesized underlying population. One form of this relationship may be formulated as follows:

$$\chi^2 = \sum_{\substack{all \\ cells}} \frac{(\text{observed} - \text{expected})^2}{\text{expected}}. \qquad (1)$$

This measures the normalized squared difference between an image (observed) and the illuminant models (expected). In terms of measuring the difference between image histogram h and a model histogram m, the measure can be reformulated as a bin-by-bin operation as follows:

$$\chi^2(h,m) = \sum_i \frac{(h_i - \mu_i)^2}{\mu_i}, \text{ with } \mu_i = \frac{h_i + m_i}{2}.$$

Here h_i and μ_i are respectively the measured image histogram value and the expected value for the i'th bin, where the expected value is defined simply as the average of the corresponding

image and model values. After substituting for μ_i and eliminating constants, the expression to be minimized over all models is simply:

$$\min : \chi^2 = \sum_i \frac{(h_i - m_i)^2}{h_i + m_i}. \tag{2}$$

[32] This match metric is computed for each model, and the illuminant that produces the minimum value of this function is taken to be the best match illuminant for the image data. In terms of equation (1), it is the illuminant that minimizes the difference between the observed values in the image and the values expected under the illuminant. In rough statistical terms, as the value of this χ^2 metric decreases, the percentage of data supporting the hypothesis that the corresponding model is the true illuminant of the scene increases.

[33] Figure 3 shows example match surfaces for a given image and the illuminant grid of Figure 1 under different match metrics (correlation, L1, L2, and chi-squared). The index of the best match illuminant is also shown for each. Notice that correlation and L2 produce the identical selection , whereas L1 and chi-squared result in slightly different choices.

[34] Algorithms of embodiments of the present invention may be implemented in software on a general-purpose computer or on a special-purpose computing device such as a DSP. Embodiments may also be implemented by dedicated circuitry such as an ASIC. Embodiment processes may be implemented in any image processing pipeline that outputs an image for display, for retrieval, for indexing or for other purposes.

[35] What is claimed is: